

RECEIVED

APR 01 2002

TECH CENTER 1600/2900

0300

1635

OIPE

#9/R.T.
4/24

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,963

DATE: 12/18/2001

TIME: 16:41:38

Raw Seq.
Listing

Input Set : N:\Crif3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

1 <110> APPLICANT: MUSC FOUNDATION FOR RESEARCH DEVELOPMENT
 2 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSIS AND TREATMENT
 3 OF CANCER BASED ON THE TRANSCRIPTION FACTOR ETS2
 4 <130> FILE REFERENCE: 9175-006-228
 5 <140> CURRENT APPLICATION NUMBER: 09/841,963
 6 <141> CURRENT FILING DATE: 2001-04-25
 7 <150> PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/27805
 W--> 8 <151> PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-23
 9 <160> NUMBER OF SEQ ID NOS: 34
 10 <170> SOFTWARE: FastSEQ for Windows Version 3.0
 12 <210> SEQ ID NO: 1
 13 <211> LENGTH: 2269
 14 <212> TYPE: DNA
 15 <213> ORGANISM: Homo sapiens
 16 <220> FEATURE:
 17 <221> NAME/KEY: CDS
 18 <222> LOCATION: (292)...(1698)
 19 <400> SEQUENCE: 1

20	ccgtttcctc ccctcccctc cactcggcgc tccctccttc ctctccctc ctccctcctc	60
21	ctccccgtcc tgaagagcgc gccgcgtggg ggacggcccg gttacttctt ccagagactg	120
22	acgagtgcgg tgctcgtcca gctcagagct ccgggagccg cccggccagc gtccggcctc	180
23	cctgatcgtc tctggccggc gccctcgccc tcgcccggcg cgcaccgagc agccgcgggc	240
24	gccgagcagc caccgtcccg accaagcgcc ggccctgccc gcagcggcag g atg aat	297
25		Met Asn
26		1
27	gat ttc gga atc aag aat atg gac cag gta gcc cct gtg gct aac agt	345
28	Asp Phe Gly Ile Lys Asn Met Asp Gln Val Ala Pro Val Ala Asn Ser	
29	5 10 15	
30	tac aga ggg aca ctc aag cgc cag cca gcc ttt gac acc ttt gat ggg	393
31	Tyr Arg Gly Thr Leu Lys Arg Gln Pro Ala Phe Asp Thr Phe Asp Gly	
32	20 25 30	
33	tcc ctg ttt gct gtt ttt cct tct cta aat gaa gag caa aca ctg caa	441
34	Ser Leu Phe Ala Val Phe Pro Ser Leu Asn Glu Glu Gln Thr Leu Gln	
35	35 40 45 50	
36	gaa gtg cca aca ggc ttg gat tcc att tct cat gac tcc gcc aac tgt	489
37	Glu Val Pro Thr Gly Leu Asp Ser Ile Ser His Asp Ser Ala Asn Cys	
38	55 60 65	
39	gaa ttg cct ttg tta acc ccg tgc agc aag gct gtg atg agt caa gcc	537
40	Glu Leu Pro Leu Leu Thr Pro Cys Ser Lys Ala Val Met Ser Gln Ala	
41	70 75 80	
42	tta aaa gct acc ttc agt ggc ttc aaa aag gaa cag cgg cgc ctg ggc	585
43	Leu Lys Ala Thr Phe Ser Gly Phe Lys Lys Glu Gln Arg Arg Leu Gly	
44	85 90 95	
45	att cca aag aac ccc tgg ctg tgg agt gag caa cag gta tgc cag tgg	633
46	Ile Pro Lys Asn Pro Trp Leu Trp Ser Glu Gln Gln Val Cys Gln Trp	
47	100 105 110	
48	ctt ctc tgg gcc acc aat gag ttc agt ctg gtg aac gtg aat ctg cag	681

ENTERED

RAW SEQUENCE LISTING

DATE: 12/18/2001

PATENT APPLICATION: US/09/841,963

TIME: 16:41:38

Input Set : N:\Crf3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

49	Leu Leu Trp Ala Thr Asn Glu Phe Ser Leu Val Asn Val Asn Leu Gln	
50	115 120 125 130	
51	agg ttc ggc atg aat ggc cag atg ctg tgt aac ctt ggc aag gaa cgc	729
52	Arg Phe Gly Met Asn Gly Gln Met Leu Cys Asn Leu Gly Lys Glu Arg	
53	135 140 145	
54	ttt ctg gag ctg gca cct gac ttt gtg ggt gac att ctc tgg gaa cat	777
55	Phe Leu Glu Leu Ala Pro Asp Phe Val Gly Asp Ile Leu Trp Glu His	
56	150 155 160	
57	ctg gag caa atg atc aaa gaa aac caa gaa aag aca gaa gat caa tat	825
58	Leu Glu Gln Met Ile Lys Glu Asn Gln Glu Lys Thr Glu Asp Gln Tyr	
59	165 170 175	
60	gaa gaa aat tca cac ctc acc tcc gtt cct cat tgg att aac agc aat	873
61	Glu Glu Asn Ser His Leu Thr Ser Val Pro His Trp Ile Asn Ser Asn	
62	180 185 190	
63	aca tta ggt ttt ggc aca gag cag gcg ccc tat gga atg cag aca cag	921
64	Thr Leu Gly Phe Gly Thr Glu Gln Ala Pro Tyr Gly Met Gln Thr Gln	
65	195 200 205 210	
66	aat tac ccc aaa ggc ggc ctc ctg gac agc atg tgt ccg gcc tcc aca	969
67	Asn Tyr Pro Lys Gly Gly Leu Leu Asp Ser Met Cys Pro Ala Ser Thr	
68	215 220 225	
69	ccc agc gta ctc agc tct gag cag gag ttt cag atg ttc ccc aag tct	1017
70	Pro Ser Val Leu Ser Ser Glu Gln Glu Phe Gln Met Phe Pro Lys Ser	
71	230 235 240	
72	cgg ctc agc tcc gtc agc gtc acc tac tgc tct gtc agt cag gac ttc	1065
73	Arg Leu Ser Ser Val Ser Val Thr Tyr Cys Ser Val Ser Gln Asp Phe	
74	245 250 255	
75	cca ggc agc aac ttg aat ttg ctc acc aac aat tct ggg act ccc aaa	1113
76	Pro Gly Ser Asn Leu Asn Leu Leu Thr Asn Asn Ser Gly Thr Pro Lys	
77	260 265 270	
78	gac cac gac tcc cct gag aac ggt gcg gac agc ttc gag agc tca gac	1161
79	Asp His Asp Ser Pro Glu Asn Gly Ala Asp Ser Phe Glu Ser Ser Asp	
80	275 280 285 290	
81	tcc ctc ctc cag tcc tgg aac agc cag tcg tcc ttg ctg gat gtg caa	1209
82	Ser Leu Leu Gln Ser Trp Asn Ser Gln Ser Ser Leu Leu Asp Val Gln	
83	295 300 305	
84	cgg gtt cct tcc ttc gag agc ttc gaa gat gac tgc agc cag tct ctc	1257
85	Arg Val Pro Ser Phe Glu Ser Phe Glu Asp Asp Cys Ser Gln Ser Leu	
86	310 315 320	
87	tgc ctc aat aag cca acc atg tct ttc aag gat tac atc caa gag agg	1305
88	Cys Leu Asn Lys Pro Thr Met Ser Phe Lys Asp Tyr Ile Gln Glu Arg	
89	325 330 335	
90	agt gac cca gtg gag caa ggc aaa cca gtt ata cct gca gct gtg ctg	1353
91	Ser Asp Pro Val Glu Gln Gly Lys Pro Val Ile Pro Ala Ala Val Leu	
92	340 345 350	
93	gcc ggc ttc aca gga agt gga cct att cag ctg tgg cag ttt ctc ctg	1401
94	Ala Gly Phe Thr Gly Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu	
95	355 360 365 370	
96	gag ctg cta tca gac aaa tcc tgc cag tca ttc atc agc tgg act gga	1449
97	Glu Leu Leu Ser Asp Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,963

DATE: 12/18/2001

TIME: 16:41:38

Input Set : N:\Crif3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

```

98          375          380          385
99  gac gga tgg gag ttt aag ctc gcc gac ccc gat gag gtg gcc cgc cgg      1497
100  Asp Gly Trp Glu Phe Lys Leu Ala Asp Pro Asp Glu Val Ala Arg Arg
101          390          395          400
102  tgg gga aag agg aaa aat aag ccc aag atg aac tac gag aag ctg agc      1545
103  Trp Gly Lys Arg Lys Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser
104          405          410          415
105  cgg ggc tta cgc tac tat tac gac aag aac atc atc cac aag acg tcg      1593
106  Arg Gly Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ser
107          420          425          430
108  ggg aag cgc tac gtg tac cgc ttc gtg tgc gac ctc cag aac ttg ctg      1641
109  Gly Lys Arg Tyr Val Tyr Arg Phe Val Cys Asp Leu Gln Asn Leu Leu
110  435          440          445          450
111  ggg ttc acg ccc gag gaa ctg cac gcc atc ctg ggc gtc cag ccc gac      1689
112  Gly Phe Thr Pro Glu Leu His Ala Ile Leu Gly Val Gln Pro Asp
113          455          460          465
114  acg gag gac tgaggtcgcc gggaccaccc tgagccggcc ccaggctcgt      1738
115  Thr Glu Asp
117  ggactgagtg ggaagcccat cctgaccagc tgccctccgag gacccaggaa aggcaggatt      1798
118  gaaaatgtcc aggaaagtgg ccaagaagca gtggccttat tgcaccccaa accacgcctc      1858
119  ttgaccaggc tgccctccctt gtggcagcaa cggcacagct aattctactc acagtgcctt      1918
120  taagtgaaaa tggtcgagaa agaggcaccg ggaagccgctc ctggcgcctg gcagtcctgt      1978
121  ggacgggatg gttctggctg tttgagattc tcaaaggagc gagcatgtcg tggacacaca      2038
122  cagactatatt ttagattttc ttttgccttt tgcaaccagg aacagcaaat gcaaaaactc      2098
123  tttgagaggg taggaggggtg ggaaggaaac aaccatgtca tttcagaagt tagtttgtat      2158
124  atattataat aatcttataa ttgttctcag aatcccttaa cagttgtatt taacagaaat      2218
125  tgtatattgt aattttaaata aattatataa ctgtatttga aataagaatt c      2269
127 <210> SEQ ID NO: 2
128 <211> LENGTH: 469
129 <212> TYPE: PRT
130 <213> ORGANISM: Homo sapiens
131 <400> SEQUENCE: 2
132  Met Asn Asp Phe Gly Ile Lys Asn Met Asp Gln Val Ala Pro Val Ala
133    1          5          10          15
134  Asn Ser Tyr Arg Gly Thr Leu Lys Arg Gln Pro Ala Phe Asp Thr Phe
135          20          25          30
136  Asp Gly Ser Leu Phe Ala Val Phe Pro Ser Leu Asn Glu Glu Gln Thr
137          35          40          45
138  Leu Gln Glu Val Pro Thr Gly Leu Asp Ser Ile Ser His Asp Ser Ala
139          50          55          60
140  Asn Cys Glu Leu Pro Leu Thr Pro Cys Ser Lys Ala Val Met Ser
141  65          70          75          80
142  Gln Ala Leu Lys Ala Thr Phe Ser Gly Phe Lys Lys Glu Gln Arg Arg
143          85          90          95
144  Leu Gly Ile Pro Lys Asn Pro Trp Leu Trp Ser Glu Gln Gln Val Cys
145          100          105          110
146  Gln Trp Leu Leu Trp Ala Thr Asn Glu Phe Ser Leu Val Asn Val Asn
147          115          120          125
148  Leu Gln Arg Phe Gly Met Asn Gly Gln Met Leu Cys Asn Leu Gly Lys

```

RAW SEQUENCE LISTING

DATE: 12/18/2001

PATENT APPLICATION: US/09/841,963

TIME: 16:41:38

Input Set : N:\Crf3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

```

149          130          135          140
150      Glu Arg Phe Leu Glu Leu Ala Pro Asp Phe Val Gly Asp Ile Leu Trp
151      145          150          155          160
152      Glu His Leu Glu Gln Met Ile Lys Glu Asn Gln Glu Lys Thr Glu Asp
153          165          170          175
154      Gln Tyr Glu Glu Asn Ser His Leu Thr Ser Val Pro His Trp Ile Asn
155          180          185          190
156      Ser Asn Thr Leu Gly Phe Gly Thr Glu Gln Ala Pro Tyr Gly Met Gln
157          195          200          205
158      Thr Gln Asn Tyr Pro Lys Gly Gly Leu Leu Asp Ser Met Cys Pro Ala
159      210          215          220
160      Ser Thr Pro Ser Val Leu Ser Ser Glu Gln Glu Phe Gln Met Phe Pro
161      225          230          235          240
162      Lys Ser Arg Leu Ser Ser Val Ser Val Thr Tyr Cys Ser Val Ser Gln
163          245          250          255
164      Asp Phe Pro Gly Ser Asn Leu Asn Leu Leu Thr Asn Asn Ser Gly Thr
165          260          265          270
166      Pro Lys Asp His Asp Ser Pro Glu Asn Gly Ala Asp Ser Phe Glu Ser
167          275          280          285
168      Ser Asp Ser Leu Leu Gln Ser Trp Asn Ser Gln Ser Ser Leu Leu Asp
169      290          295          300
170      Val Gln Arg Val Pro Ser Phe Glu Ser Phe Glu Asp Asp Cys Ser Gln
171      305          310          315          320
172      Ser Leu Cys Leu Asn Lys Pro Thr Met Ser Phe Lys Asp Tyr Ile Gln
173          325          330          335
174      Glu Arg Ser Asp Pro Val Glu Gln Gly Lys Pro Val Ile Pro Ala Ala
175          340          345          350
176      Val Leu Ala Gly Phe Thr Gly Ser Gly Pro Ile Gln Leu Trp Gln Phe
177          355          360          365
178      Leu Leu Glu Leu Leu Ser Asp Lys Ser Cys Gln Ser Phe Ile Ser Trp
179      370          375          380
180      Thr Gly Asp Gly Trp Glu Phe Lys Leu Ala Asp Pro Asp Glu Val Ala
181      385          390          395          400
182      Arg Arg Trp Gly Lys Arg Lys Asn Lys Pro Lys Met Asn Tyr Glu Lys
183          405          410          415
184      Leu Ser Arg Gly Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Ile His Lys
185          420          425          430
186      Thr Ser Gly Lys Arg Tyr Val Tyr Arg Phe Val Cys Asp Leu Gln Asn
187          435          440          445
188      Leu Leu Gly Phe Thr Pro Glu Glu Leu His Ala Ile Leu Gly Val Gln
189      450          455          460
190      Pro Asp Thr Glu Asp
191      465
193 <210> SEQ ID NO: 3
194 <211> LENGTH: 84
195 <212> TYPE: PRT
196 <213> ORGANISM: Homo sapiens
197 <400> SEQUENCE: 3
198      Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,963

DATE: 12/18/2001

TIME: 16:41:38

Input Set : N:\Crif3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

```

199      1          5          10          15
200      Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
201              20          25          30
202      Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
203              35          40          45
204      Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
205              50          55          60
206      Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
207      65              70          75          80
208      Tyr Arg Phe Val
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 84
212 <212> TYPE: PRT
213 <213> ORGANISM: Mus musculus
214 <400> SEQUENCE: 4
215      Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
216      1          5          10          15
217      Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
218              20          25          30
219      Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
220              35          40          45
221      Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
222              50          55          60
223      Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
224      65              70          75          80
225      Tyr Arg Phe Val
227 <210> SEQ ID NO: 5
228 <211> LENGTH: 84
229 <212> TYPE: PRT
230 <213> ORGANISM: Gallus gallus
231 <400> SEQUENCE: 5
232      Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
233      1          5          10          15
234      Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
235              20          25          30
236      Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
237              35          40          45
238      Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
239              50          55          60
240      Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
241      65              70          75          80
242      Tyr Arg Phe Val
244 <210> SEQ ID NO: 6
245 <211> LENGTH: 84
246 <212> TYPE: PRT
247 <213> ORGANISM: Xenopus laevis
248 <400> SEQUENCE: 6
249      Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
250      1          5          10          15

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/841,963

DATE: 12/18/2001

TIME: 16:41:39

Input Set : N:\Crf3\RULE60\09841963.raw

Output Set: N:\CRF3\12182001\I841963.raw

L:8 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD